Impact of Camera Defects on Detection in Autonomous Driving

Tomasz Frelek
The Ohio State University
Columbus, OH, USA
frelek.1@osu.edu

ABSTRACT

Camera degradation poses a serious risk to the reliability of autonomous vehicle perception systems. This study examines the impact of common camera defects on object detection performance using the YOLOv11 model and the Udacity Self-Driving Car Dataset. Simulated defects caused significant drops in detection accuracy metrics, particularly for smaller objects. We show that adversarial training on defect-augmented images greatly improves robustness without compromising baseline accuracy. These results highlight the need for defect-resilient models to ensure long-term autonomous driving safety.

KEYWORDS

autonomous vehicles, object detection, camera degradation, adversarial training, computer vision

ACM Reference Format:
Tomasz Frelek. 2025. Impact of Camera Defects on Detection in Autonomous Driving. In *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems (SenSys '25), November 2025*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Autonomous driving systems rely heavily on camera-based object detection to interpret and navigate complex environments. These vision systems are critical for recognizing other vehicles, pedestrians, road signs, lanes, and various obstacles, forming the foundation of decision-making processes in self-driving cars. In fact, some companies, like Tesla[15], rely primarily on their camera suite to guide their autonomous vehicles. However, cameras, as with any hardware, are susceptible to a range of defects, both due to manufacturing error and adverse conditions in the field. Over time, these defects can impair image quality, leading to degraded detection performance.

As autonomous vehicles become increasingly prevalent on public roads, the aging of their hardware components, particularly camera sensors, presents a growing concern. Unlike mechanical parts that are regularly maintained or replaced, the calibration/maintinence of camera sensor suites is unregulated, and they often remain in use well beyond their initial calibration period [1]. Over time, exposure to environmental stressors such as heat, vibration, and radiation can degrade these sensors, leading to accumulation of defects like dead pixels, increased noise, or diminished focus. As fleets of autonomous vehicles age, the likelihood of such defects increases, potentially compromising the performance of object detection systems that are vital for safe navigation.

The consequences of such impairments pose significant safety risks. Misidentifying or failing to detect objects can result in collisions, endangering passengers and pedestrians alike. In safety-critical applications like autonomous driving, even small sensor degradations can have disproportionate impacts on system reliability. Despite rapid increase of deployment of autonomous (and semi-autonomous) vehicles on the road, there has been little corresponding research exploring the effects of camera defects on object detection performance. As autonomous vehicle use is projected to increase[2], addressing this gap is vital to ensure the robustness and safety of autonomous vehicles.

1.1 Camera Defects

Cameras used in autonomous vehicles are sensitive devices that must operate reliably in a wide range of conditions. However, like any complex hardware, these sensors are susceptible to both physical and digital defects arising from rough field conditions. Autonomous vehicles operate in dynamic environments, facing constant exposure to vibration, extreme temperatures, moisture, dust, and radiation, all of which can negatively affect camera performance. Such conditions can significantly accelerate sensor degradation over time. As autonomous vehicles are expected to function for extended periods without frequent hardware replacement, understanding and mitigating the risks associated with both inherent and environmental camera defects is essential to maintaining reliable object detection and ensuring the safety of the vehicle and its surroundings.

Cameras in autonomous vehicles are subject to a range of defects that can broadly be categorized as physical or digital. Physical defects arise from the hardware components themselves, including issues like lens scratches, focus drift, sensor misalignment, or lens contamination from dirt, moisture, or debris. These physical imperfections directly impair the quality of the visual data captured, distorting or obscuring objects in the environment. Over time, exposure to harsh conditions such as vibration, extreme temperatures, or weather can exacerbate these defects, degrading sensor performance.

In contrast, digital defects stem from the internal processing of image data. These include faulty pixels/scanlines, sensor noise, RBG sensor drift, etc. While physical defects alter the actual scene captured by the camera, digital defects affect how that scene is represented in the data, introducing artifacts that may not be visible to the naked eye but can disrupt object detection algorithms.

SenSys '25, November 2025, Location TBD

ACM ISBN 978-1-4503-XXXX-X/25/11
https://doi.org/10.1145/1122445.1122456

Both defect types can accumulate over time, and without proper monitoring or correction, they can compromise the safety and reliability of autonomous systems. Understanding effects of physical and digital defects is essential for designing robust perception systems that maintain performance across the lifespan of the vehicle.

2 RELATED WORK

The reliability of camera systems in autonomous vehicles is a relatively new area of study, without much literature on the topic. To our knowledge, no long-term studies on vehicular-grade cameras are public, likely due to the proprietary nature of the technology compounded by the long time-frame necessary to conduct such a survey. However, the long-term reliability of consumer camera hardware was analyzed in a study by SquareTrade [9], which reported failure rates of approximately 11% over two years, with 6.6% due to malfunctions. While this study focused on consumer-grade cameras, the findings are relevant to autonomous vehicles, which often operate with similar CMOS-based sensors. This suggests that hardware degradation over time is a realistic concern.

Some research has been dedicated to studying the effects of image defects on object detection in autonomous systems. Filo et al. [7] contributed valuable insights by benchmarking the robustness of 3D object detection models against common corruptions such as noise, blur, and weather-induced artifacts. Their work provides an important foundation for understanding how perception systems perform under less-than-ideal conditions. However, their study primarily focused on synthetic perturbations, such as Gaussian noise or image shearing added in post-production, rather than defects cameras experience via real-world degradation. As such, questions surrounding long-term sensor reliability remain largely unaddressed.

In terms of real-world defects, research on blemish artifacts has also been a point of study. In particular, Augustine et al. [3] demonstrated that even minor optical obstructions can disrupt object detection models. Their work underscores the importance of unblemished optical pathways for accurate perception. However, their work focuses on manufacturing defects as opposed to defects that arise naturally from overuse and degradation.

3 DESIGN

To our knowledge, there are no studies that analyze the effects of camera defects resulting from natural wear-and-tear on autonomous driving scenarios. Toward that end, we have designed an experiment to systematically evaluate how various physical and digital defects affect detection performance. By digitally simulating these defects at different severity levels and applying them to a standard object detection algorithm, we assess their influence on model accuracy and reliability. This approach allows us to determine potential failure points that may arise as vehicle cameras age or encounter harsh operating conditions.

3.1 Dataset

This study utilizes the Udacity Self-Driving Car Dataset (2020)[13], a publicly available dataset designed to support research in autonomous vehicle perception. The dataset includes images captured

Figure 1: An example of some of the diverse conditions captured in the Udacity dataset.

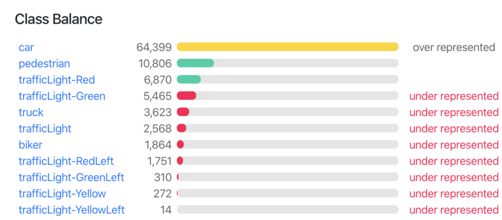

Figure 2: The label distribution of the Udacity dataset is severly skewed, over-representing a few key classes while under-representing many others.

from a vehicle-mounted sensor suite in various driving environments, including urban streets, highways, and suburban roads. The data was collected in the San Fransisco Bay area in 2016-2017. It encompasses 97,942 labels across 11 classes and 15,000 images. This dataset was chosen because it provides a suite of diverse driving scenarios, while maintaining a manageable size (13GB unzipped) and a simple annotation format. The dataset is not perfect, as it features an over-representation of some label classes while severely under-representing others, although this is an issue faced by many other datasets, and is not a problem unique to the Udacity dataset.

3.2 Model

For this study, we utilize YOLOv11 [17], the latest evolution in the YOLO family of object detection models from Ultralytics[16]. YOLOv11 represents a significant advancement over previous iterations, offering enhanced speed, accuracy, and efficiency over the previous mainstream versions. The YOLOv11 version introduces several architectural innovations, notably a C3K2 (an upgrade to Cross Stage Partial blocks) backbone, SPFF (Spatial Pyramid Pooling Fast) modules, and C2PSA (Cross Stage Partial with Spatial Attention) attention mechanisms, which collectively enhance the model's ability to detect small and partially occluded objects while maintaining fast inference speeds.

In this study, YOLOv11 serves as the object detection model that we will use evaluate the impact of camera defects. Its cutting-edge architecture makes it an ideal candidate for testing robustness under degraded visual conditions, providing insights into how modern

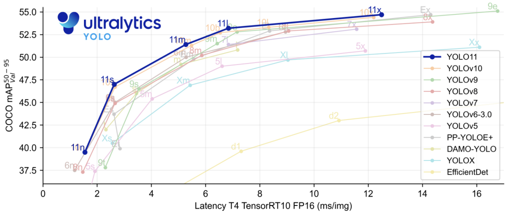

Figure 3: A comparison of YOLOv11 performance to the performance of previous YOLO models.

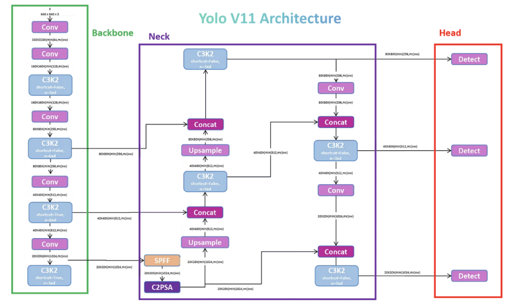

Figure 4: An overview of the architecture of YOLOv11.

perception systems withstand the challenges posed by aging or compromised sensors in autonomous vehicles.

3.3 Defects

In this study, we focus on a diverse set of camera defects that can arise from both physical damage and digital sensor degradation. These defects are representative of real-world scenarios where camera performance deteriorates over time due to environmental exposure, wear, and internal sensor aging. Understanding their individual and combined effects on object detection performance is essential for ensuring the robustness and safety of autonomous vehicle systems. The following defect types are considered:

- **Cracks (Physical Defect):** Cracks in the camera lens or housing typically result from mechanical impacts, such as debris striking the camera, minor collisions, or even stress fractures caused by repeated vibrations or thermal expansion and contraction over time. Such defects distort or partially block the camera's field of view, affecting object recognition.
- **Mud/Dirt (Physical Defect):** Mud, dirt, dust, or water droplets accumulating on the lens surface lead to partial occlusion of the image. Such contamination is common in outdoor environments, especially in autonomous vehicles exposed to varied road and weather conditions, and can obscure critical visual information.
- **Chromatic Aberration (Physical Defect):** Chromatic aberration is caused by lens imperfections, where different wavelengths of light are not focused at the same point [4]. This results in color fringing around high-contrast edges. While modern lenses mitigate this through design, cheaper optics, lens wear, or temperature-induced expansion can worsen this effect over time.
- **Tint (Digital/Physical Defect):** Tinting refers to an unwanted color shift in captured images. Physically, this can occur due to wear or degradation of anti-reflective coatings on the lens, or from long-term exposure to sunlight (UV damage)[12]. Digitally, tinting may arise from sensor aging or miscalibration [5] of color balance over time, especially as sensors operate under varying lighting conditions. Tinting can impair color-based object recognition tasks, such as detecting traffic lights or interpreting road signs.
- **Broken Scanlines (Digital Defect):** Broken scanlines manifest as horizontal or vertical lines across the image due to sensor readout failures or transmission errors. These defects are particularly associated with CMOS (Complementary Metal-Oxide-Semiconductor) sensors, which are commonly used in automotive cameras due to their low power consumption, high-speed readout, and integration capabilities[6]. CMOS sensors read out pixel data row by row, so when errors occur in this sequential process, they can manifest as broken or missing scanlines in the final image, which can then affect the detections of objects in frame.
- **Sensor Blooming (Digital Defect):** Sensor blooming occurs when bright light sources (e.g., sunlight, headlights) saturate the sensor's pixels, causing the excess charge to spill into adjacent pixels. Most modern cameras have in-built suppresion systems to prevent/mitigate sensor blooming [11], however, such mechanisms can degrade due to sensor aging or by exposure to high-intensity lighting, which damages the pixel structure.

By simulating these defects at varying severity levels, this study aims to systematically evaluate their individual and combined impacts on the performance of object detection models, specifically YOLOv11. The inclusion of both physical artifacts and digital sensor errors provides a comprehensive assessment of camera degradation scenarios relevant to real-world autonomous driving environments.

4 EVALUATION

This section outlines the experimental design, computational environment, defect simulation process, and the resulting performance evaluation of the object detection model under various camera degradation scenarios. By systematically introducing both physical and digital defects into the dataset, we assess how these imperfections affect the detection capabilities of the YOLOv11 model and explore potential mitigation strategies.

4.1 Computational Resources

All experiments were conducted using Google Colab [8], which provided access to cloud-based GPU acceleration for efficient model training and evaluation. Specifically, the hardware configuration included NVIDIA Tesla T4 GPUs, Intel Xeon CPUs, and 16 GB of system memory, a setup well-suited for testing object detection models in a scalable, resource-constrained environment. The Ultralytics YOLOv11 implementation, built on the PyTorch framework, was used for both training and inference. This combination of tools enabled real-time performance evaluations while taking advantage of YOLOv11's advanced architecture optimizations. The Google Colab environment also provided pre-installed support for CUDA 11.x and Python 3.9+, streamlining the experimental workflow.

4.2 Model Training

For this study, we employed a YOLOv11 model, specifically YOLO11s, a lightweight version of the model with 9.4M parameters [14]. The model was initialized using pre-trained weights from the COCO dataset, providing a strong baseline for general object detection. To adapt the model to the specific domain of autonomous driving, we fine-tuned it on the Udacity Self-Driving Car Dataset, which includes labeled objects such as vehicles, pedestrians, and traffic signs. The dataset was unmodified at this stage, meaning it contained only the original images, as the defects were digitally introduced later in the experimental sequence.

The training process was conducted on Google Colab, leveraging NVIDIA Tesla T4 GPUs. Key hyperparameters included an image size of 512x512 pixels, batch size of 16, and 10 epochs of training, and the Adam optimizer with an initial learning rate of 0.001. We used the pre-built Ultralytics pipeline to train the model. Our model took a little over an hour to train, and we were unable to train for longer due to colab's usage limits (about 2 hours per session). After training, the best-performing model checkpoint (based on validation mAP) was selected for evaluation under defect conditions.

4.3 Evaluation Metrics

The primary metric used to assess the object detection performance in this study is the mean Average Precision at an IoU threshold of 0.5 (mAP@50). This metric evaluates the balance between precision (the proportion of correct positive detections) and recall (the proportion of actual objects correctly detected) at a standard Intersection over Union (IoU) threshold of 0.5. mAP@50 provides a straightforward indication of the model's ability to accurately localize and classify objects under various defect conditions.

In addition to mAP@50, the Ultralytics YOLOv11 framework automatically generates several other informative evaluation metrics:

- **Precision**: Measures the proportion of predicted positive detections that are correct, indicating the reliability of the model when it makes detections.
- **Recall**: Measures the proportion of actual objects that are correctly detected, reflecting the model's completeness in identifying all relevant objects.

While mAP@50 serves as the primary benchmark for comparing the model's robustness under different defect scenarios, the additional metrics provide finer-grained insights. These metrics help identify specific performance trade-offs, such as whether the model tends to miss objects (low recall) or generate false positives (low precision) under certain types of camera degradation.

4.4 Digital Modification

To evaluate the impact of various digital camera defects on object detection performance, we applied simulated imperfections to the test images. These modifications replicate both physical defects and sensor-level digital artifacts. The following defects were generated:

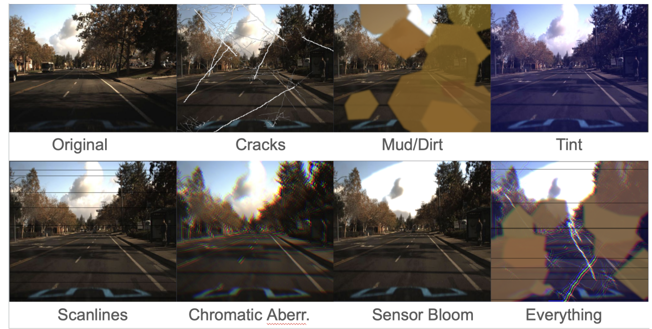

Figure 5: An example of how different defects affect an image.

- **Mud/Dirt Splatter**: Mud splatters were generated by overlaying irregularly shaped brown polygons onto the image, simulating dirt or mud on the camera lens. The shapes were randomly positioned and blurred for realism, with adjustable opacity to simulate different severities.
- **Cracks**: Cracks were overlaid using pre-generated alpha-blended crack images, resized to match the base image. These cracks introduced irregular line patterns across the image to mimic physical damage to the lens or protective housing.
- **Tinting**: A color tint was applied by blending a solid color (red, green, or blue) over the entire image. The intensity of the tint was adjustable, simulating lens discoloration or sensor calibration drift that leads to color bias.
- **Broken Scanlines**: Horizontal black lines were randomly drawn across the image to simulate broken scanlines. These lines mimic sensor readout failures or electronic interference, which result in missing rows of pixel data in a CMOS system.
- **Chromatic Aberration**: Chromatic aberration was simulated by shifting the red and blue color channels in opposite directions while keeping the green channel fixed. This introduces color fringing around object edges, mimicking optical imperfections in the lens.
- **Sensor Blooming**: Sensor blooming was replicated by detecting bright regions in the image and applying a Gaussian blur to those areas, causing the brightness to "bleed" into adjacent pixels. This simulates the charge overflow in image sensors when exposed to high-intensity light sources like headlights or sunlight.

4.5 Results

To analyze the efficacy of the model on object detection under various defects, we will rely on the mAP50, recall, and precision statistic of the car, pedestrian, and trafficLight-Red classes, as the other classes suffer from underepresentation in the dataset, which corresponds to worse baseline accuracy for said classes.

4.5.1 ***No Modification***. To establish a baseline accuracy, we evaluated the YOLOv11 model on the original test set devoid of modifications.

As shown in Table 1, the model demonstrates strong performance for cars, with high precision (0.847), recall (0.788), and mAP@50 (0.843), reflecting reliable detection due to the large number of

Table 1: Detection Metrics for Unmodified Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.847	0.788	0.843
pedestrian	744	2253	0.749	0.532	0.629
trafficLight-Red	518	1441	0.918	0.696	0.803

[a]P: Precision for bounding boxes; R: Recall.

Table 2: Detection Metrics for Crack Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.753	0.743	0.784
pedestrian	744	2253	0.627	0.457	0.506
trafficLight-Red	518	1441	0.823	0.624	0.709

car instances in the dataset. TrafficLight-Red also achieves high precision (0.918), ensuring minimal false positives, though recall (0.696) suggest occasional missed detections and some localization imprecision, likely due to the small object size. Pedestrian detection lags with lower recall (0.532) and mAP@50 (0.629), indicating frequent missed detections, perhaps attributable to the small frame and varied poses of many pedestrian cases.

4.5.2 ***Crack Modification***.

To test the effects of cracks on object detection, each image in the test set was modified to contain linear cracks with varying length, branching, and thickness.

As shown in Table 2, car detection remains robust, with precision (0.753), recall (0.743), and mAP@50 (0.784), suggesting that linear occlusions from cracks have minimal impact on larger, well-structured objects. Pedestrian detection shows moderate degradation, with precision (0.627), recall (0.457), and mAP@50 (0.506), indicating that cracks disrupt the detection of smaller objects but less severely than broader occlusions like mud. TrafficLight-Red detection remains strong, with precision (0.823), recall (0.624), and mAP@50 (0.709), reflecting that while cracks interfere with localized regions, the model can still effectively detect small but well-defined objects like traffic lights. Overall, cracks introduce narrow, localized occlusions, which tend to affect small or thin objects more than large ones but still allow relatively high detection performance across most classes.

4.5.3 ***Mud/Dirt Modification***.

To test the effects of mud occlusions on object detection, each image in the test set was modified to contain an mud splatter of varying size, shape, and opacity (as described in the previous section).

As shown in Table 3, in the presence of mud occlusions, car detection maintains a relatively stable precision (0.749) though recall and mAP50 fall to 0.547 and 0.589 respectively indicating a sizable drop in detection accuracy. Pedestrian detection is even more affected, with lower precision (0.616), recall (0.360), and mAP@50 (0.394). TrafficLight-Red retains the highest precision (0.856), but recall (0.437) and mAP@50 (0.526) show notable degradation, reflecting difficulty detecting small, partially covered objects. We posit that this performance drop across all classes is due to mud splatters occluding critical features, such as edges and textures, that the model relies on for detection. Smaller objects, like pedestrians and traffic lights, are especially vulnerable since even minor occlusions can cover significant portions of their bounding boxes, leading to missed detections.

Table 3: Detection Metrics for Mud Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.749	0.547	0.589
pedestrian	744	2253	0.616	0.360	0.394
trafficLight-Red	518	1441	0.856	0.437	0.526

Table 4: Detection Metrics for Chromatic Aberration Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.813	0.566	0.678
pedestrian	744	2253	0.486	0.333	0.351
trafficLight-Red	518	1441	0.614	0.371	0.412

Table 5: Detection Metrics for Tinted Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.832	0.632	0.708
pedestrian	744	2253	0.585	0.385	0.413
trafficLight-Red	518	1441	0.843	0.557	0.656

4.5.4 ***Chromatic Aberration Modification***.

To test the effects of image chromatic aberration on object detection, each image in the test set had its red and blue channels shifted by a random amount, to simulate chromatic aberration to various degrees of severity.

As shown in Table 4,chromatic aberration caused car detection to fall, with precision (0.813), recall (0.566), and mAP@50 (0.678), showing that even larger objects affected by color fringing. Pedestrian detection, experiences a significant decline, with precision (0.486), recall (0.333), and mAP@50 (0.351), indicating that misaligned color channels disrupt the model's ability to localize and classify smaller, detail-dependent objects. TrafficLight-Red also degrades, with precision (0.614), recall (0.371), and mAP@50 (0.412), reflecting challenges in detecting color-dependent objects when channel shifts alter key color cues.

4.5.5 ***Tint Modification***.

To test the effects of image tinting on object detection, each image in the test set was modified to contain red, green, or blue tint at a varying intensity (as described in the previous section).

As shown in Table 5, in the presence of tinting, car detection remains strong, with precision (0.832), recall (0.632), and mAP@50 (0.708), indicating relatively minimal impact from color shifts on larger, well-defined objects. Pedestrian detection shows reduced performance, with precision at 0.585, recall at 0.385, and mAP@50 at 0.413, reflecting moderate sensitivity to tint-induced color distortions. TrafficLight-Red achieves high precision (0.843) and reasonable recall (0.557), with mAP@50 at 0.656, suggesting the model maintains robustness for detecting red traffic signals despite color bias.

Table 6: Detection Metrics for Broken Scanline Test Set

Class	Images	Instances	Box(P)	R	mAP50
car	2585	12853	0.779	0.764	0.805
pedestrian	744	2253	0.648	0.431	0.496
trafficLight-Red	518	1441	0.877	0.597	0.721

Table 7: Detection Metrics for Sensor Bloom Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.843	0.786	0.839
pedestrian	744	2253	0.728	0.541	0.620
trafficLight-Red	518	1441	0.906	0.664	0.778

4.5.6 ***Scanline Modification.*** To test the effects of broken scanlines on object detection, each image in the test set was modified to contain a varied amount of scanlines blacked out.

As shown in Table 6, in the presence of broken scanlines, car detection remains effective, with precision (0.779), recall (0.764), and mAP@50 (0.805), indicating that horizontal scanline disruptions minimally impact large objects with strong structural features. Pedestrian detection shows moderate degradation, with precision (0.648), recall (0.431), and mAP@50 (0.496), suggesting that partial occlusions along scanlines interfere with recognizing smaller, detailed objects. TrafficLight-Red detection stays strong, with precision (0.877), recall (0.597), and mAP@50 (0.721), reflecting robustness against localized interruptions, though performance still dips compared to baseline. Overall, broken scanlines introduce localized occlusion artifacts, which are less disruptive for large or well-defined objects but may be more impactful on smaller, detail-reliant classes like pedestrians.

4.5.7 ***Blooming Modification.*** To instigate the effects of sensor blooming on object detection, each image in the test set was modified to bloom at a random intensity, as described in the previous section.

As seen in Table 7, sensor blooming seems to have minimal impact on object detection, and all three classes' metrics stay close to their baselines. We posit that this is because typically only the brightest sections of an image are susceptible to bloom, which during the day (when this dataset was collected) is almost always the sky, meaning that blooming doesn't cause any substantial interference with detecting objects located on the ground. However, the results may differ if the data were collected at night, as then headlights and streetlights would be affected by bloom, theoretically lowering the detection accuracy.

4.5.8 ***All Modifications.*** To analyze the effects that multiple defects would have, we modified each image in the test set to display a random series of defects. Each defect had a 50% chance to be included in a given test image, making for a completely randomly modified test set.

As shown in Table 8, the presence of multiple defects degrades detection performance the most severly for all classes. Performance for cars degrades notably, with precision dropping to 0.665, recall to 0.365, and mAP@50 falling to 0.411, indicating a significant loss

Table 8: Detection Metrics for Fully Modified Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.665	0.365	0.411
pedestrian	744	2253	0.385	0.174	0.164
trafficLight-Red	518	1441	0.698	0.267	0.321

Figure 6: An example of how different defects affect object detection in an image. The left image in each pair is the ground truth, the right image in each pair is the model's prediction under the effects of defects.

in detection accuracy compared to baseline conditions. Pedestrian detection suffers even more severely, with precision at 0.385, recall at 0.174, and mAP@50 at 0.164, reflecting both poor localization and high miss rates. TrafficLight-Red detection also weakens, though it retains relatively higher precision (0.698) and mAP@50 (0.321) compared to pedestrians, but recall is reduced to 0.267, indicating that defects particularly hinder the model's ability to detect and localize smaller objects consistently.

Seemingly, the effects of multiple defects compound on an image and reduce the model's accuracy far more than just a single defect.

4.6 Remedies

There have been some attempts to combat the defects cameras may experience. For example, Parente et al. [10] devised a machine learning algorithm that detects cracks in an image, allowing for a system of swift identification that informs the camera's user that defects are present. In a similar vain, Wu et al. [18] developed a system that does the same for dirt on a camera.

These systems, while important safeguards, fail to address the core problem: the degradation of object detection performance itself. While detecting the presence of a defect is valuable for maintenance and diagnostics, it does not prevent or mitigate the immediate loss of perception capabilities in real-time operations. In high-stakes environments like autonomous driving, the system must remain resilient to visual impairments even before maintenance actions can be taken. Thus, it is equally critical to design object detection models that are robust to visual distortions, ensuring reliable performance despite camera degradation.

To that end, we propose a system of adversarial training that will teach the detection model to be robust to potential defects encountered in the field. Adversarial training is a common machine learning technique [19]; in this context, adversarial training

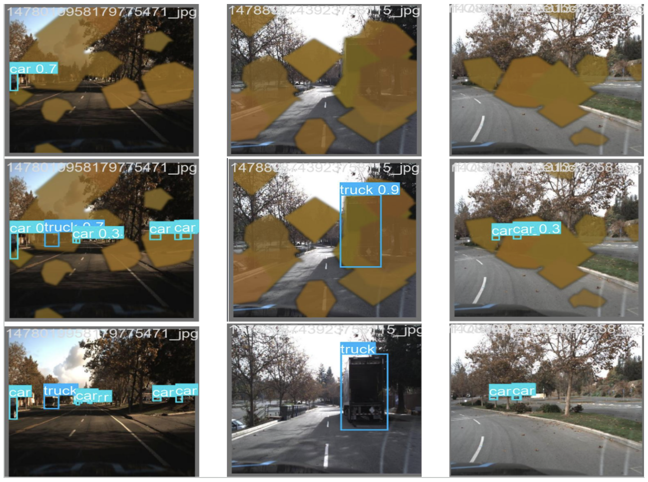

Figure 7: An example of the improvements of the adversarially trained model. The images in the top row are the predictions of the baseline model, the middle row are the predictions of the adversarially trained model, the bottom row are ground truths.

involves exposing the model to defect-augmented images during the training process. By systematically introducing variations such as mud splatters, cracks, chromatic aberration, broken scanlines, and sensor blooming into the training dataset, the model learns to recognize and correctly classify objects under a wider range of degraded visual conditions. This forces the model to rely on more invariant and robust features rather than superficial texture cues that may be disrupted by camera defects.

To demonstrate this, we modified 50% of the images in the Udacity training set to contain mud splatters, and then trained a second YOLO11s model on the modified training set. This model was trained with the same hyperparameters and hardware restrictions as the original model.

This second model showed great generalizabilty, as seen in Tables 9 and 10. When evaluated on the original unmodified test set, the adversarially trained model performed almost identically to the baseline model in all metrics, meaning it maintains accuracy in ideal conditions. Additionally, when evaluated on the test set modified to contain mud splatters (the exact same one used to evaluate the unmodified model), the adversarially trained model demonstrated remarkable resilience, maintaining mAP50, precision, and recall scores that were very close to those of the unmodified test set (for all three main classes no less). On the mud test set, the adversarially trained model exhibited a 24.5%, 16%, 24.2% improvement in mAP50 scores for cars, pedestrians, and red traffic lights, respectively.

These results clearly demonstrate that adversarial training provides a viable pathway to harden object detection models against camera degradation without sacrificing baseline performance in ideal conditions. We chose to adversarially train on the mud defect as an illustrative example, since it was the defect that caused the highest loss in accuracy. However, the model could very easily be adversarially trained against any of the defects presented in this paper.

Table 9: Detection Metrics for Adversarially Trained Model on Unmodified Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.844	0.779	0.834
pedestrian	744	2253	0.739	0.490	0.596
trafficLight-Red	518	1441	0.896	0.696	0.798

Table 10: Detection Metrics for Adversarially Trained Model on Mud Test Set

Class	Images	Instances	P	R	mAP50
car	2585	12853	0.818	0.761	0.814
pedestrian	744	2253	0.676	0.459	0.544
trafficLight-Red	518	1441	0.874	0.666	0.768

As autonomous vehicles are deployed in increasingly diverse and challenging environments, equipping perception systems with such resilience will be crucial for maintaining safety, reliability, and operational continuity over the long term.

5 CHALLENGES

During the course of this study, several challenges were encountered, spanning both technical limitations and experimental complexities.

One primary challenge involved the accurate simulation of physical camera defects in a digital environment. While digital modifications such as cracks, dirt overlays, tinting, and chromatic aberration filters were applied to mimic real-world degradation, these simulations may not fully capture the complex optical distortions that occur in actual damaged lenses. Replicating non-linear effects, such as light scattering or multi-angle reflections caused by physical lens imperfections, remains difficult without hardware-in-the-loop testing. This limits the generalizability of the findings to some extent, as real-world degradation could produce more complex artifacts.

Another challenge was hardware limitations. While YOLOv11 is optimized for speed and efficiency, our Google Colab sessions presented practical limitations, particularly with session timeouts and GPU availability fluctuations, which sometimes impacted experiment continuity. Additionally, given greater hardware resources, we could have attempted to train a larger model, potentially YOLOv11l (Large) or even YOLOv11x (Extra-large), and for many more epochs.

Additionally, quantifying the severity of defects posed a subjective challenge. There is no standardized metric for calibrating digital defect intensities against real-world camera degradation. This made it difficult to precisely correlate digital defect severity with actual sensor wear or environmental damage encountered in autonomous vehicles, and ultimetly, we had to make educated guesses as to how camera defects manifest themselves.

6 CONCLUSION

In this study, we systematically evaluated the impact of various real-world camera defects on object detection performance in autonomous driving scenarios. By simulating a range of physical and digital imperfections, we demonstrated that even moderate levels of degradation can significantly impair the ability of state-of-the-art models like YOLOv11 to reliably detect objects. Small and detail-dependent classes, such as pedestrians and traffic lights, were found to be especially vulnerable to defects, while larger objects like cars exhibited greater resilience under certain conditions.

Importantly, our results show that the cumulative effects of multiple concurrent defects can cause severe degradation, far exceeding the impact of isolated issues. These findings highlight the need for proactive strategies to enhance the robustness of perception systems against sensor aging and environmental wear.

We also explored the use of adversarial training as a mitigation strategy, showing that models trained with defect-augmented data can substantially recover lost performance under defect conditions without sacrificing baseline accuracy on clean images. This approach offers a promising pathway to harden autonomous vehicle perception systems against real-world operational challenges.

Future work should focus on developing more realistic defect simulation pipelines, investigating hardware-in-the-loop evaluations, and expanding adversarial training frameworks to cover a broader range of degradation scenarios. As autonomous vehicles continue to proliferate, ensuring the long-term reliability of their perception systems will be critical for maintaining public safety and trust.

REFERENCES

[1] 2024. New San Antonio Company Targets Untapped, Unregulated Vehicle Industry. [Online]. Available: https://www.mysanantonio.com/business/article/vehicle-tech-industry-texas-20020558.php.

[2] 2024. Semi-autonomous Vehicle Market Size, Share & Trends Analysis Report. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/semi-autonomous-vehicle-market.

[3] Nitin Augustine, Maximilian Schwab, Steffen Klarmann, Christian Pfefferer, and Alexander Schiendorfer. 2024. Impact of Blemish Artifacts on Object Detection Models. *Procedia Computer Science* 223 (2024), 61–68. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050924000619.

[4] M. G. Bouma and A. J. Goldsmith. 2023. Lens Abnormalities. In *StatPearls*. StatPearls Publishing, Treasure Island (FL). [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK597386/.

[5] R. Clark. 2022. Sensor Calibration and Color. [Online]. Available: https://clarkvision.com/articles/sensor-calibration-and-color/.

[6] EE Times Asia. 2020. Teardown: Tesla's Hardware Retrofits for Model 3. [Online]. Available: https://www.eetasia.com/teslas-hardware-retrofits-for-model-3/.

[7] A. Filo, B. Laugraud, M. Sanchez, R. Martín-Martín, D. Scaramuzza, A. Gaidon, and Y. Zhou. 2023. Benchmarking Robustness of 3D Object Detection to Common Corruptions. *arXiv preprint arXiv:2303.11040* (2023).

[8] Google. 2024. Google Colaboratory. [Online]. Available: https://colab.research.google.com/.

[9] SquareTrade Inc. 2009. Camera Failure Rates: New Data on How Consumer Cameras Fail. [Online]. Available: https://www.squaretrade.com/htm/pdf/Camera_failure_study.pdf.

[10] Luigi Parente, Eugenia Falvo, Cristina Castagnetti, Francesca Grassi, Francesco Mancini, Paolo Rossi, and Alessandro Capra. 2022. Image-Based Monitoring of Cracks: Effectiveness Analysis of an Open-Source Machine Learning-Assisted Procedure. *PubMed* (2022). [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC8876482/.

[11] Hamamatsu Photonics. [n. d.]. CCD Saturation and Blooming. [Online]. Available: https://hamamatsu.magnet.fsu.edu/articles/ccdsatandblooming.html.

[12] J. Raabe, A. Kini, and A. G. Lee. 2019. Optical Lens Tinting—A Review of its Functional Mechanism, Efficacy, and Applications. *European Ophthalmic Review* 12, 2 (2019), 85–87. [Online]. Available: https://touchophthalmology.com/retina-vitreous/journal-articles/optical-lens-tinting-a-review-of-its-functional-mechanism-efficacy-and54-applications/.

[13] Roboflow. 2020. Self-Driving Car Dataset. [Online]. Available: https://public.roboflow.com/object-detection/self-driving-car.

[14] Roboflow. 2025. YOLOv11s vs YOLOv11n Model Comparison. [Online]. Available: https://roboflow.com/compare-model-sizes/yolo11s-vs-yolo11n.

[15] Tesla, Inc. 2024. Autopilot. [Online]. Available: https://www.tesla.com/en_AE/autopilot.

[16] Ultralytics. 2024. Ultralytics. [Online]. Available: https://www.ultralytics.com.

[17] Ultralytics. 2024. YOLOv11 Models Documentation. [Online]. Available: https://docs.ultralytics.com/models/yolo11/.

[18] Zhijun Wu, Yuchang Wang, Lujia Ran, Zongjie Hu, Zhuolin Liu, Xiangrui Ding, Zhengzhe Chen, Yuxin Shi, and Zhengyuan Gao. 2023. Research on Sensor Dirt Recognition of Autonomous Vehicle Camera Based on Deep Learning. In *Proceedings of the Second International Conference on Informatics, Networking, and Computing (ICINC 2023)*. [Online]. Available: https://proceedings.spiedigitallibrary.org/conference-proceedings-of-spie/13078/130780M/Research-on-sensor-dirt-recognition-of-autonomous-vehicle-camera-based/10.1117/12.3024700.full.

[19] M. Zhao, L. Zhang, J. Ye, H. Lu, B. Yin, and X. Wang. 2024. Adversarial Training: A Survey. *arXiv preprint arXiv:2410.15042* (2024). [Online]. Available: https://arxiv.org/abs/2410.15042.